# Assignment

Since 2013, the Citi Bike Program has implemented a robust infrastructure for collecting data on the program's utilization. Through the team's efforts, each month bike data is collected, organized, and made public on the [Citi Bike Data] (<https://www.citibikenyc.com/system-data>) webpage.

Based on data received, please analyze the data. Analysis of some data should display your capabilities of data analytics.

# Data

**However, often there is too much data to “export” and some data is lost.**

Graphical user interface, application

Description automatically generated

This data has been processed to remove trips that are taken by staff as they service and inspect the system, trips that are taken to/from any of our “test” stations (which we were using more in June and July 2013), and any trips that were below 60 seconds in length (potentially false starts or users trying to re-dock a bike to ensure it's secure).

Please be aware of your software program’s row limitations as you are viewing the data. Many of the CSV files contain more than 1 million rows. After downloading, you will need to use a large data tool / visualizer (like Tableau, Alteryx, R, or others) to view and analyze the full data sets.

Data has converted to excel and combined by monthly tabs to represent full year of 2020. The data has been cleaned with Data Interpreter of Tableau. Within spreadsheet the gender numbers were changed to represent gender in alpha format instead of number (1 = male, 2 = female, 0 = unknown). An age based column was also inserted into monthly tabs based on end of 2020, since month’s were not provided for birth.

# Analysis

Graphical user interface, text, application

Description automatically generated  
<https://public.tableau.com/app/profile/mark8269/viz/CitiBike2020_EmerBootcamp/Dashboard4?publish=yes>